Constraints on Analogical Mapping: The Effects of Similarity & Order

Mark T. Keane

Department of Computer Science,
Trinity College, University of Dublin,
Dublin 2, IRELAND
mkeane@cs.tcd.ie

Tim Ledgeway & Stuart Duff

School of Psychology,
University of Wales College of Cardiff,
PO Box 901, Cardiff, WALES
duff@cardiff.ac.uk

Abstract

One of the central problems in analogical mapping is overcoming the ambiguities which can occur when matching up corresponding concepts in two domains of knowledge; specifically, to ensure that one-to-many and many-to-one matches are resolved to be one-to-one matches. Various analogy theories have attempted to deal with these problems by maintaining that analogical matching is constrained in various ways. For example, that only predicates of the same structural-type are matched, that primacy is given to matches that are similar or identical, and that a match which comes before an alternative match is preferred. Two experiments are reported, involving an attribute-mapping problem, which isolate the effects of similarity and order. The first shows that the semantic similarity of predicates in the two domains has a facilitating effect on analogical mapping when other constraints are held constant. The second experiment shows that analogical mapping is sensitive to the order in which matches are made. The implications of these results for current computational models of analogy are discussed, with a special emphasis on the consequences that order effects have for connectionist models.

Introduction

When we draw an analogy the hard work is done by the process that performs an analogical mapping between two domains of knowledge (see e.g., Burstein 1986; Carbonell 1986; Gentner 1983, 1989; Gick & Holyoak 1980, 1983; Holyoak 1985; Keane 1985, 1988a). Ranged about this core process are the other processes of analogical thinking; for example, processes concerned with the representation of the domains and the retrieval of analogues from long-term memory. Typically, when an analogical mapping is made two distinct computations occur; first, the corresponding concepts in both domains are *matched* and, second, a portion of the conceptual structure of one domain is *transferred* (or carried over) into the other domain to form the basis of analogical inferences. For example, if you want to understand why electrons revolve around the nucleus in the atom, and are told that the atom is like a miniature solar system (see Gentner 1983), you would *match* the corresponding REVOLVES relations in both domains and *transfer* relations of ATTRACTION from the solar system domain to apply in the atom domain.

In general, people appear to match up the concepts in two domains with an effortlessness that masks some tricky computational problems. Even though two domains may have a one-to-one correspondence between their parts, they can hold a great potential for ambiguous matches; for one-to-many and many-to-one matches between their parts. Psychologically, these problems only surface when we make the mapping task difficult, such as in the following problem from Holyoak & Thagard (1989):

A	B
Bill is smart.	Fido is hungry.
Bill is tall.	Blackie is friendly.
Tom is timid.	Blackie is frisky.
Tom is tall.	Rover is hungry.
Steve is smart.	Rover is friendly.

In this problem, subjects are asked to say which things in list A correspond to which things in list B (ignoring the meaning of the words). Essentially, subjects have to discover a one-to-one mapping between all the individuals and attributes in list A and list B. This is quite a difficult task as a lot of ambiguous matches have to be resolved. For example, *smart* may match *hungry* or *friendly* or *frisky* and the correct match can only be determined by eliminating the inconsistent matches which follow from all but one of these matches. The unique one-to-one mapping which solves the problem involves matching Steve and Fido, Bill and Rover, Tom and Blackie, smart and hungry, tall and friendly and timid and frisky.

Palmer (1989) has pointed out that any adequate theory of analogical mapping, which accounts for problems like the above one, will have to operate at several levels of description (for similar ideas see Marr 1982). At the highest level, one needs to characterise the *informational constraints* implied by the task situation; this level is concerned with describing what an analogy is, what needs to be computed to produce appropriate outputs given certain inputs (like Marr's computational level). Below this level is the level of *behavioural constraints* which have to capture the empirical facts of people's observable analogical behaviour (Marr's algorithmic level). Hence, this level should include constraints that predict when one analogy is harder than another and the sorts of errors that people produce. Finally, there is the level of hardware constraints which aims to capture the hardware primitives of analogical thought (Marr's hardware level).

Current theories of analogy propose many informational constraints but less attention have been given to the behavioural and hardware constraints. In this paper we test one of the well-established, information constraints, and a prediction from behavioural constraints previously proposed by Keane (1990a).

Informational Constraints on Analogical Mapping

We have already seen, that there are several difficulties associated with achieving optimal analogical mappings. Three informational constraints have been proposed to solve these problems (see e.g., Gentner 1983; Holyoak & Thagard 1989; Keane 1990a).

The most important set of constraints are *structural constraints*. These constraints are used to enforce a one-to-one mapping between the two domains (Falkenhainer, Forbus & Gentner 1986; Holyoak & Thagard 1989). This is done using several techniques:

- *matches are made only between entities of the same type*; for example, only attributes are matched with attributes, objects with objects and two-place predicates with two-place predicates. This reduces the total number of matches that needs to be considered (see Gentner 1983; Holyoak & Thagard 1989).
- *exploit structural consistency*; that is, if the propositions REVOLVES(A B) and REVOLVES(C D) match, then the arguments of both should also be matched appropriately, A with C and B with D. This is especially useful in eliminating many-to-one and one-to-many matches (see Falkenhainer et al. 1986, 1989).
- *favour systematic sets of matches* (Gentner's 1983, systematicity principle); this proposes that if one has two alternative sets of matches then the match-set with the most higher-order connectivity should be chosen. This aids the choice of an optimal match-set from among many match-sets.

These techniques have been shown to be very powerful. In many cases, structural constraints alone can find the optimal mapping between two domains (as in the above attribute-mapping case).

A *similarity constraint* can also be used to reduce the number of matches considered or to disambiguate between alternative matches. When this constraint is applied only identical concepts are matched between the two domains (Gentner 1983) or, more loosely, semantically-similar concepts are matched (Gick & Holyoak 1980). Semantic similarity can be used to disambiguate matches because if one match in a set of one-to-many matches is more similar than the others then it can be preferred.

A final constraint on analogical mapping is the *pragmatic constraint* (e.g., Holyoak 1985; Keane 1985). Again, this may disambiguate a set of matches. For example, if in a certain analogical mapping situation one match is pragmatically more important (or goal-relevant) than other alternatives then it will be preferred over these alternatives.

These informational constraints constitute a very, high-level specification of what makes a particular comparison between two domains an analogical comparison. As such, they can and, indeed, have been implemented by many different algorithms. Falkenhainer et al.'s (1986, 1989) *Structure Mapping Engine* (SME) implements structural and similarity constraints in a serial fashion that generates all the possible match-sets between two domains and then chooses the best match-set according to structural criteria. Forbus & Oblinger (1990) have extended SME to implement the pragmatic constraint. Holyoak & Thagard's (1989) *Analogical Constraint Mapping Engine* (ACME) uses parallel constraint satisfaction to ease a network of possible matches into a single stable set of matches, in accordance with the three constraints. Keane & Brayshaw's (1988; Keane 1991) *Incremental Analogy Machine* (IAM) implements the same constraints, using a different serial algorithm to SME that computes parts of the analogy incrementally.

Behavioural Constraints on Analogical Mapping

Any model which solely implements the above informational constraints can make certain behavioural predictions. For example, Skorstadt, Falkenhainer & Gentner (1987) have shown that SME can produce outputs that parallel broad trends in subjects' soundness ratings of different comparisons. Holyoak & Thagard (1989) have shown that the number of cycles their network goes through before settling into a correct mapping bears some correspondence to the difficulty subjects find in analogical mapping. However, because these models do not embody behavioural constraints, they fail to capture significant facets of human analogical performance. While they may predict the outputs of analogical mapping they do not predict response-time differences for different analogies (see Experiment 2). They also tend to be silent on the source and nature of errors in analogical performance.

Keane (1988b, 1990a, 1990b) has elaborated two behavioural constraints; first, that analogising should be subject to constraints imposed by *working memory limitations* and, second, that differential *background knowledge about the target domain* affects the ease with which analogical transfers are validated. It should be noted that both of these constraints are predicted to affect response times even though the content output from analogising may be unchanged. Hence, the separation between informational and behavioural constraints. Working memory limitations may also result in errors (see Keane 1990b).

Keane (1990a) has also argued that working memory limitations have wider implications for the nature of analogical processing. In particular, that human analogical processes will reduce the computational overheads in analogising. As such, human analogisers will tend to only map portions of the base domain (see Keane 1985) or will choose one optimal mapping. This is the theoretical rationale for the Incremental Analogy Machine. Rather than generating all possible mappings, IAM tries to build just one mapping resolving ambiguities in a local fashion. It may even use fairly arbitrary means to resolve an ambiguity. For example, if a one-to-many set of matches is encountered and there are no grounds for selecting one match over the others, IAM will simply choose the first match in the set.

Clearly, there will be occasions when IAM will commit itself to one mapping only to find that it is non-optimal. When this happens it will undo this mapping and start constructing an alternative mapping. So, while IAM tries to avoid constructing all possible mappings it may backtrack repeatedly to construct several alternative mappings.

As we shall see in Experiment 2, IAM predicts that the order in which the attributes are presented in versions of the attribute-mapping problem will differentially affect subjects' response times. This prediction is to be contrasted with the outputs of the SME and ACME models. Both of these models do not predict such order effects.

The Effects of Similarity and Order on Analogical Mapping

Many experiments have shown that varying the similarity of two analogues affects the ease of analogical mapping (e.g., Keane 1985; Gentner & Landers 1985; Gick & Holyoak 1980; Holyoak & Koh 1987). It should also be possible to show this in a very precise fashion in the above attribute-mapping problem.

In Experiment 1, we systematically vary the number of predicates that are similar while controlling the other constraints. For instance, a modified version of the problem with similar attributes should be much easier:

A	B
Bill is intelligent.	Fido is clever.
Bill is tall.	Blackie is big.
Tom is timid.	Blackie is shy.
Tom is tall.	Rover is clever.
Steve is intelligent.	Rover is big.

In Experiment 2, we examine the prediction of IAM that ordering effects should be observed under some conditions in this attribute-mapping problem. Holyoak & Thagard (1989) had randomised the order of presentation of the attributes in both lists thus abolishing any order effects. To appreciate how these effects might arise consider what makes the problem so difficult.

In the abstract version of the attribute-mapping problem, each list has two individuals (e.g., Bill and Tom) with two attributes and a remaining individual (i.e., Steve) who has just one attribute. This is very important because matching up the single individuals in both lists (i.e., Steve and Fido) is the key to achieving the isomorphic mapping. The presence of these single individuals with one attribute (which we will call *singletons*) disambiguates the set of matches between the two lists (this also applies to the single attribute in both lists). We expect, on the basis of simulations using IAM, that order has an effect on matching and that people match items incrementally starting with the first item in list A. Hence, the problem should be easier when the singletons are placed at the beginning of the list where they can be matched first.

Experiment 1: Similarity Effects

To test the prediction that semantic similarity can facilitate analogical mapping three groups received three different versions of the attribute-mapping problem which had either no similar attributes, one set of similar attributes or three sets of similar attributes. We predicted that subjects would produce the correct one-to-one mapping more rapidly as a function of the increasing similarity between the two domains.

Method

Subjects & Design. Twenty-four undergraduates at the University of Wales College of Cardiff took part voluntarily in the experiment. The experiment used a between-subject design and subjects were assigned randomly to one of the three conditions. Three subjects were dropped from the experiment before data analysis because they misunderstood the experimental instructions. Data analysis was carried out on the remaining 21 subjects, who were equally distributed across the three conditions.

Materials. We used three versions of the attribute-mapping problem (see Table 1). Each version had two lists of attributes. In each list, there were three individuals and three attributes; two individuals had two attributes and one individual had a single attribute. The three versions differed in terms of the number of attributes that were similar in both lists. In the None-Similar version none of the attributes were semantically similar; in the One-Similar version one set of the attributes was semantically similar ("intelligent" in list A and "clever" in list B); in the All-Similar version, all the attributes in one list had a semantically-similar parallel attribute in the second list (see Table 1).

A	*B (None-Similar)*
Bill is intelligent.	Fido is hungry.
Bill is tall.	Blackie is friendly.
Tom is timid.	Blackie is frisky.
Tom is tall.	Rover is hungry.
Steve is intelligent.	Rover is friendly.
B (One-Similar)	*B (All-Similar)*
Fido is clever.	Fido is clever.
Blackie is friendly.	Blackie is big.
Blackie is frisky.	Blackie is shy.
Rover is clever.	Rover is clever.
Rover is friendly.	Rover is big.

Table 1 The Different Problems Used in the Three Conditions of Experiment 1 (Only B differed in each)

Subjects were instructed in writing that their "task is to figure out what in the left set corresponds to what in the right set of sentences". A single column below list A listed the names of the individuals and attributes in that list. Next to each was a space for subjects to write the corresponding name or attribute from list B. The order of sentences in each list was randomised with the proviso that sentences with attributes about the same individual were kept together.

Procedure. Subjects were first shown the instructions which they were asked to read carefully. They were then shown the problem and asked to solve it. A stop-watch was used to time them from this point to when they solved the problem. If subjects produced an incorrect answer they were told so and asked to continue solving the problem. Only when the correct answer was produced was the clock stopped and the elapsed time recorded.

Results & Discussion

The results corroborated our expectations. The presence of semantic similarities between the elements of the two domains has an important facilitating effect on the ease of analogical mapping (see Figure 1). The elapsed time taken to solve the problem gradually decreases across the three conditions, with the None-Similar Condition being the slowest (M = 210.9 secs.), the One-Similar Condition being faster (M = 164.9 secs.) and the All-Similar Condition being the fastest (M = 69.7 secs.) [$F(2, 18) = 8.747$, $p < .005$]. This is the first demonstration that domains which systematically-differ in terms of their similarity, give rise to systematically-differing solution times.

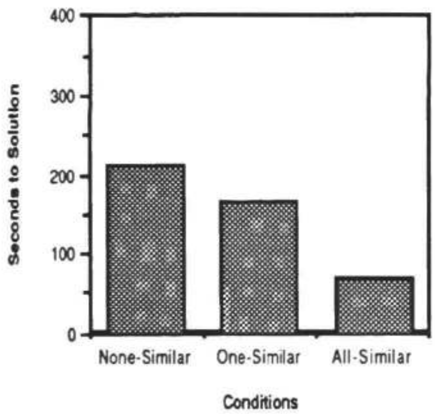

Figure 1 The Mean Solution Times in the Conditions of Experiment 1

Experiment 2: Order Effects

Our second experiment tested whether there were order effects on analogical mapping, as is predicted by the Incremental Analogy Machine. As we said earlier, the key to finding the isomorphism, in the abstract version of the problem, lies in matching the two singletons in both lists. So, if one group receives the singletons first they should match them and solve the problem quicker, than if they form the singleton match later. In two models of analogy, SME and ACME, simple changes of this sort do not significantly change the course or outputs of analogical mapping. However, IAM's processing is radically different for both versions. When IAM matches the singletons first it does not have to undo its matches and so finds the isomorphism faster than when the singletons come later in the lists. Specifically, IAM does not backtrack at all in the Singleton-First case but has to backtrack four times in the Singleton-Last case.

Method

Subjects & Design. Twenty-three undergraduates at the University of Wales College of Cardiff took part voluntarily in the experiment. As before, the experiment had a between-subject design and subjects were assigned at random to one of the two conditions. Again, three subjects had to be excluded from the experiment prior to data analysis because they misunderstood the experimental instructions. Data analysis was carried out on the remaining 20 subjects, who were equally distributed across the two conditions.

Materials & Procedure. The materials consisted of two abstract versions of the attribute-matching problem (see Table 2). In the Singleton-First version the singletons were at the top of both lists, while in the Singleton-Last version the singleton in list A was in the last position, while the Singleton in list B was in the first position (see Table 2). The order of the remaining sentences was randomised as before.

The remainder of the materials were as in Experiment 1, with the exception of the additional sentence: "The meaning of the words in the sentences is irrelevant". Holyoak and Thagard used this sentence when they gave subjects the abstract version of the problem.

Subjects were shown the sheet containing the instructions and problem, and were timed, as before.

Singleton-First	
A	B
Steve is smart.*	Fido is hungry.*
Bill is tall.	Blackie is friendly.
Bill is smart.	Blackie is frisky.
Tom is tall.	Rover is hungry.
Tom is timid.	Rover is friendly.
Singleton-Last	
A	B
Bill is smart.	Fido is hungry.*
Bill is tall.	Blackie is friendly.
Tom is timid.	Blackie is frisky.
Tom is tall.	Rover is hungry.
Steve is smart.*	Rover is friendly.

Table 2 The Two Versions of the Problem Used in Experiment 2 (*the singleton)

Results & Discussion

The slight change in the ordering of the singletons has a marked effect on the ease of analogical mapping (see Figure 2). Subjects in the Singleton-First condition were almost twice as fast at solving the problem (M = 178.0 secs) compared to the Singleton-Last condition (M = 363.1 secs) [Mann-Whitney $U = 7$, $p < .005$, 1-tailed].

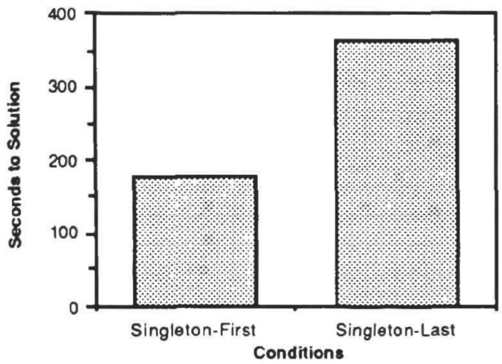

Figure 2 The Mean Solution Times for the Conditions in Experiment 2

This result clearly shows that the order in which the attributes are presented affects subjects' response-times. In this experiment, we have looked at a situation in which the 'easy' problem had the two singletons in the *same position*. IAM does not predict that being in the same position *per se* is important. The singletons could be in different positions and still show different response times. For example, if the singleton in list A was first and that in list B was last then IAM would predict a slower response time than, say, a problem where the list B singleton was third. These are predictions that need to be examined in future experiments.

Conclusions

Both experiments provide strong evidence for the effects of similarity and order on analogical mapping. If the predicates in both domains bear some unambiguous similarity to one another then the mapping is considerably facilitated. Furthermore, these predicates need not be identical to aid matching but may be merely similar (e.g., big and tall).

We have also seen that under certain conditions the order in which the information in a domain is presented can have significant effects on the ease of analogical mapping. As we have seen these results are not predicted by the SME and ACME models. IAM does, however, predict these effects because it has mechanisms that try to reduce the computational overheads on analogical mapping, by heuristically disambiguating one from a set of matches using order, in accordance with the behavioural constraint of working memory limitations.

The present work was deeper implications for connectionist models in cognitive science. ACME was a important program because it showed that an example of high-level cognition, namely analogical thought, could be modelled by connectionist techniques. The present work suggests that ACME is not a wholly adequate model of analogical behaviour. It fails to capture the more serial aspects of analogising manifested in the present study. This, therefore, keeps open the debate in this area on whether the nature of high-level cognition can be captured singularly by connectionist, 'brain-style' AI.

References

Burstein, M.H. 1986. Analogical learning with multiple models. In T.M. Mitchell, J.G. Carbonell & R.S. Michalski (Eds.), *Machine Learning: A Guide to Current Research*. Lancaster: Kluwer Academic Publishers.

Carbonell, J.G. 1986. Derivational analogy: A theory of reconstructive problem solving and expertise acquisition. In R.S. Michalski, J.G. Carbonell & T.M. Mitchell (Eds.), *Machine Learning II: An Artificial Intelligence Approach* . Los Altos, Calif: Kaufmann.

Falkenhainer, B., Forbus, K.D., and Gentner, D. 1986. Structure-mapping engine. Proceedings of the Annual Conference of the American Association for Artificial Intelligence. Los Altos: Kaufmann.

Falkenhainer, B., Forbus, K.D., & Gentner, D. 1989. Structure-mapping engine. *Artificial Intelligence* 41:1-63.

Forbus, K.D. & Oblinger, D. 1990. Making SME greedy and pragmatic. Twelfth Annual Conference of the Cognitive Science Society. Hillsdale: Erlbaum.

Gentner, D. 1983. Structure-mapping: A theoretical framework for analogy. *Cognitive Science* 7 :155-170.

Gentner, D. 1989. Mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and Analogical Reasoning*. Cambridge: CUP.

Gentner, D. & Landers, R. 1985. Analogical reminding: A good match is hard to find. Proceedings of the International Conference on Systems, Man and Cybernetics , Tucson, Arizona, November.

Gick, M.L., & Holyoak, K.J. 1980. Analogical problem solving. *Cognitive Psychology* 12:306-355.

Gick, M.L., & Holyoak, K.J. 1983. Schema induction in analogical transfer. *Cognitive Psychology* 15:1-38.

Holyoak, K.J. 1985. The pragmatics of analogical transfer. *The Psychology of Learning and Motivation* 19:59-87.

Holyoak, K.J., & Koh, K. 1987. Surface and structural similarity in analogical transfer. *Memory & Cognition* 15:332-340.

Holyoak, K.J., & Thagard, P. 1989. Analogical mapping by constraint satisfaction. *Cognitive Science* 13:295-355.

Keane, M. 1985. On drawing analogies when solving problems: A theory and test of solution generation in an analogical problem solving task. *British Journal of Psychology* 76:449-458.

Keane, M.T. 1988a. *Analogical Problem Solving*. Chichester: Ellis Horwood (New York: Wiley).

Keane, M.T. 1988b. Analogical Mechanisms. *Artificial Intelligence Review* 2(4):229-251.

Keane, M.T. 1990a. Incremental analogising: Theory and model. In K.Gilhooly et al. (Eds.). *Lines of Thinking*. Chichester: John Wiley.

Keane, M.T. 1990b. Anchoring Analogies in Past Experience. Unpublished ms.

Keane, M.T. 1991. IAM V2: A Re-Implementation in Lisp with Further Additions. CS-TCD Technical Report, May.

Keane, M.T., & Brayshaw, M. 1988. The Incremental Analogical Machine: A computational model of analogy. In D. Sleeman (Ed.), *European Working Session on Machine Learning*. London: Pitman.

Marr, D. 1982. *Vision*. San Francisco: Freeman.

Palmer, S.E. 1989. Levels of description in information processing theories of analogy. In S. Vosniadou & A. Ortony (Eds.), *Similarity and Analogical Reasoning*. Cambridge: CUP.

Skorstadt, J.; Falkenhainer, B.; and Gentner, D. 1987. Analogical processing: A simulation and empirical corroboration. Proceeding of AAAI'87. Las Altos: Morgan Kaufmann.